CONVERGENCE RATES FOR APPROXIMATIONS OF FUNCTIONALS OF SDES

RAINER AVIKAINEN

ABSTRACT. We consider upper bounds for the approximation error $\mathbb{E}|g(X) - g(\hat{X})|^p$, where X and $\hat{X}$ are random variables such that $\hat{X}$ is an approximation of X in the L_p-norm, and the function g belongs to certain function classes, which contain e.g. functions of bounded variation. We apply the results to the approximations of a solution of a stochastic differential equation at time T by the Euler and Milstein schemes. For the Euler scheme we provide also a lower bound.

1. INTRODUCTION

Convergence schemes for the solutions of SDEs are rather well known. Let X be the solution of the one-dimensional equation

$$dX_t = \sigma(t, X_t)\, dW_t + b(t, X_t)\, dt,\ X_0 = x_0,$$

where W is a standard Brownian motion, $t \in [0, T]$, and σ and b satisfy certain assumptions specified in Section 2. P.E. Kloeden and E. Platen [12] have showed that any order of strong convergence can be achieved by the strong Itô-Taylor approximations, i.e. for any order $\gamma > 0$ there exists a scheme X^π corresponding to a partition π of the interval $[0, T]$ with mesh size $|\pi|$ such that

$$\mathbb{E}\left(\sup_{0 \leq t \leq T} |X_t - X_t^\pi|\right) < C\, |\pi|^\gamma.$$

The simplest and most common examples are the Euler scheme X^E and the Milstein scheme X^M, which have the order of strong convergence 0.5 and 1, respectively.

Lately this topic has been considered by N. Hofmann, T. Müller-Gronbach and K. Ritter in [7, 8], Hofmann and Müller-Gronbach in [9], and Müller-Gronbach in [14, 15]. In these papers they cover errors with respect to both global and pointwise error criterions. The latest result concerning the pointwise error is due to Müller-Gronbach [15], where the author defines certain classes of convergence schemes and finds optimal (adaptive) schemes for each class.

The pointwise weak error $\mathbb{E}f(X_T) - \mathbb{E}f(X_T^\pi)$ is also considered by several authors, e.g. Kloeden and Platen [12], under certain smoothness conditions on f. In the case of the Euler scheme the smoothness conditions were relaxed by V. Bally and D. Talay in [2], where f was required to be only measurable and bounded. This was done under a Hörmander type condition for the infinitesimal generator of X. A recent contribution to this field is given by Moon et al. in [16].

This paper considers the error of the type $\mathbb{E}|g(X_T) - g(X_T^\pi)|^p$. The question is motivated by discretization schemes for BSDEs. The terminal condition $g(X_T)$ is approximated by $g(X_T^\pi)$, and L_p-estimates for the difference $g(X_T) - g(X_T^\pi)$ are required. If g is Lipschitz, this error returns immediately to the error of the underlying scheme. Therefore the aim of this paper is to give results for relevant non-Lipschitz functions. It is proved that if approximations $(X_t^\pi)_{t\in[0,T]}$ satisfy

$$\|X_T - X_T^\pi\|_p \leq C_p^1 \, |\pi|^\gamma$$

for some $\gamma > 0$ and all $1 \leq p < \infty$, then

$$\mathbb{E}|g(X_T) - g(X_T^\pi)|^p \leq C_p^2 \, |\pi|^{\gamma-\varepsilon}$$

for any $0 < \varepsilon < \gamma$ and for any g in a special class of functions. This class contains functions of bounded variation, polynomials and jump functions controlled by the tail distributions of X_T and X_T^π, and therefore by Minkowski's inequality the result is true for any sum of these three types of functions.

Finally, for the Euler scheme a lower bound is presented indicating that the error under consideration can not converge faster than $|\pi|^{1/2}$. This is done by giving an example, namely the geometric Brownian motion, for which the lower bound is obtained. There is still a difference, although arbitrarily small, between the upper and lower bounds, and it remains open whether the rates could be equalized. These results are achieved under certain conditions on the SDE, including the existence of a bounded density for the solution X_T.

The structure of the paper is such that the assumptions that hold throughout the paper are collected in Section 2. Sections $3-5$ contain the main results concerning upper bounds. The first result is given in Section 3, where Theorem 3.4 clarifies the convergence rate for indicator functions. This is then applied to the Euler and Milstein schemes in Theorem 3.7. The result is then extended to functions of bounded variation in Theorem 4.3 in Section 4, and applied to the Euler and Milstein schemes in Theorem 4.5. Another extension is developed in Section 5, where the result for the function class $\mathcal{G}_{p,\varphi}$ is given in Theorem 5.7. The class $\mathcal{G}_{p.\varphi}$ is then analyzed in Section 6, including the result that it contains all polynomials in Theorem 6.3. An application to the Euler and Milstein schemes is presented in Corollary 6.6. Section 7 contains a lower bound for the convergence, stated in Theorem 7.2.

Finally, a proof of a Theorem from the book of Bouleau and Lépingle [1] is presented with explicit constants in Appendix A.

2. Assumptions

We fix a terminal time $T > 0$ and suppose that $(W_t)_{t \in [0,T]}$ is a standard one-dimensional Brownian motion defined on a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t \in [0,T]})$, where the filtration is the augmentation of the natural filtration of W and $\mathcal{F} = \mathcal{F}_T$.

We consider a diffusion process X, which is a solution to

$$(2.1) \qquad \begin{cases} dX_t = \sigma(t, X_t)\, dW_t + b(t, X_t)\, dt, \\ X_0 = x_0 \end{cases}$$

with $x_0 \in \mathbb{R}$ and $\sigma, b : [0,T] \times \mathbb{R} \to \mathbb{R}$. We assume that $\sigma, b \in C^{0,1}([0,T] \times \mathbb{R})$ and for $f \in \{\sigma, b\}$ there exist constants C_T and $\alpha \geq \frac{1}{2}$ such that

  (i) $|f(t,x)| \leq C_T(1 + |x|)$,
  (ii) $|f(t,x) - f(t,y)| \leq C_T|x - y|$,
  (iii) $|f(t,x) - f(s,x)| \leq C_T(1 + |x|)|t - s|^\alpha$.

Assumptions (i)-(iii) imply the existence of a unique adapted strong solution X of the SDE (2.1), see e.g. [11, p. 289]. For estimates concerning the Milstein scheme we assume in addition that

  (iv) the state derivatives $\frac{\partial \sigma}{\partial x}$ and $\frac{\partial b}{\partial x}$ satisfy the condition (ii).

Moreover, we assume that

  (v) X_T has a bounded density.

Remark 2.1. Assumption (v) is satisfied (see [6, p. 263]), if we assume that $\sigma, b \in C_b^\infty([0,T] \times \mathbb{R})$ and σ satisfies the uniform ellipticity condition, i.e. there exists a constant β such that

$$\sigma(t,x) \geq \beta > 0 \text{ for all } (t,x) \in [0,T] \times \mathbb{R}.$$

Another sufficient condition is given by Caballero et al. in [5, Theorem 2]. They assume that σ and b are C^2 in x, the second derivatives have polynomial growth, the functions $|\sigma(0,x)|$, $|\sigma_x(t,x)|$, $|b(0,x)|$ and $|b_x(t,x)|$ are bounded, and

$$\mathbb{E}\left(\left|\int_0^t \sigma(s, X_s)^2\, ds\right|^{-p_0/2}\right) < \infty$$

for some $p_0 > 2$ and for all $t \in (0,T]$. Then there exists a continuous density f_{X_t} of X_t such that for all $p > 1$

$$f_{X_t}(x) \leq C_p \left\|\left(\int_0^t \sigma(s, X_s)^2\, ds\right)^{-1/2}\right\|_p$$

for some constant $C_p > 0$.

Denote by π a partition $0 = t_0 < t_1 < \ldots < t_n = T$ of the interval $[0, T]$, and let
$$|\pi| = \max_{0 \le i < n} |t_{i+1} - t_i|$$
be the mesh size of π. Moreover, denote an approximation of X corresponding to π by X^π. Two such approximations are the well known Euler and Milstein schemes.

Definition 2.2 (Euler scheme). Let X^E be the Euler scheme relative to π, i.e. $X_0^E = x_0$, and for $i = 0, \ldots, n-1$,
$$X_{t_{i+1}}^E = X_{t_i}^E + b(t_i, X_{t_i}^E)(t_{i+1} - t_i) + \sigma(t_i, X_{t_i}^E)(W_{t_{i+1}} - W_{t_i}).$$
Given the values at the partition points, we also define the Euler scheme in continuous time by setting
$$X_t^E = X_{t_k}^E + \sigma(t_k, X_{t_k}^E)(W_t - W_{t_k}) + b(t_k, X_{t_k}^E)(t - t_k)$$
for $t \in (t_k, t_{k+1})$. This can be written using the integral representation

(2.2)
$$\begin{aligned}X_t^E = x_0 &+ \int_0^t \sum_{k=0}^{n-1} \sigma(t_k, X_{t_k}^E) \chi_{[t_k, t_{k+1})}(u)\, dW_u \\ &+ \int_0^t \sum_{k=0}^{n-1} b(t_k, X_{t_k}^E) \chi_{[t_k, t_{k+1})}(u)\, du\end{aligned}$$

for $t \in [0, T]$ a.s., which coincides a.s. with the discrete scheme at the partition points. In this paper we need the continuous time version only for the Euler scheme.

Definition 2.3 (Milstein scheme). Let X^M be the Milstein scheme relative to π, i.e. $X_0^M = x_0$, and and for $i = 0, \ldots, n-1$,
$$\begin{aligned}X_{t_{i+1}}^M =\, &X_{t_i}^M + b(t_i, X_{t_i}^M)(t_{i+1} - t_i) + \sigma(t_i, X_{t_i}^M)(W_{t_{i+1}} - W_{t_i}) \\ &+ \frac{1}{2}\sigma(t_i, X_{t_i}^M)\sigma_x'(t_i, X_{t_i}^M)((W_{t_{i+1}} - W_{t_i})^2 - (t_{i+1} - t_i)).\end{aligned}$$

We remind that the Euler and Milstein approximations of X_T, denoted X_T^E and X_T^M, always depend on the corresponding partition π. This is omitted from the notation for simplicity.

These assumptions hold throughout the paper.

3. Indicator Functions

3.1. **General Approximation.** Suppose that we have a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and two random variables $X, \hat{X} : \Omega \to \mathbb{R}$. Consider $\hat{X}$ to be an approximation of X in the L_p-norm. First we find an estimate for the quantity $\mathbb{E}|g(X) - g(\hat{X})|$, where g is an indicator function.

Definition 3.1. Recall the non-increasing rearrangement of a random variable X, defined by $X^* : [0, 1] \to \mathbb{R} \cup \{+\infty, -\infty\}$,
$$X^*(s) := \inf\{c \in \mathbb{R} : \mathbb{P}(X > c) \le s\}.$$

Here we use the convention that $\inf \emptyset = \infty$.

Remark 3.2. Definition 3.1 is slightly different from the standard non-increasing rearrangement as defined e.g. in [4], where the absolute value of the function X is taken. However, by analoguous arguments we can show the following properties:

(i) $X^*(1) = -\infty$, $X^*(0) = \infty$ if X is not essentially bounded and $X^*(s) \in \mathbb{R}$ for $s \in (0,1)$,

(ii) X^* is right-continuous,

(iii) X^* has the same distribution as X with respect to the Lebesgue measure on $[0,1]$,

Definition 3.3. Denote the minimal slope of the function X^* from the level K by $d_X : \mathbb{R} \to [0, \infty)$,

$$d_X(K) := \inf_{\substack{s \in [0,1] \\ s \neq \alpha(K)}} \left\{ \frac{|X^*(s) - K|}{|s - \alpha(K)|} \right\},$$

where

$$\alpha(K) = \mathbb{P}(X \geq K).$$

Theorem 3.4. *Suppose that X is a random variable. Then the following assertions hold:*

(i) *If X has a bounded density f_X, then for all $K \in \mathbb{R}$, all random variables $\hat{X}$ and all $0 < p < \infty$ we have*

$$\mathbb{E}|\chi_{[K,\infty)}(X) - \chi_{[K,\infty)}(\hat{X})| \leq 3 D_X(K)^{\frac{p}{p+1}} \left\| X - \hat{X} \right\|_p^{\frac{p}{p+1}},$$

where

$$D_X(K) := \frac{1}{d_X(K)} \in (0, \sup f_X].$$

Moreover, the power $\frac{p}{p+1}$ of the L_p-norm is optimal, i.e. if

$$(3.1) \qquad \mathbb{E}|\chi_{[K,\infty)}(X) - \chi_{[K,\infty)}(\hat{X})| \leq C(X, K, p) \left\| X - \hat{X} \right\|_p^{\frac{p}{p+1}}$$

for all random variables $\hat{X}$, then the power $\frac{p}{p+1}$ can not be replaced by a power q such that $\frac{p}{p+1} < q < \infty$.

(ii) *If there exists $p_0 > 0$ such that the formula (3.1) holds for all $p_0 \leq p < \infty$, all $K \in \mathbb{R}$ and all random variables $\hat{X}$, and there exists $B_X > 0$ such that $C(X, K, p) \leq B_X$, then X has a bounded density.*

Proof. Let us first show (i). Fix $K \in \mathbb{R}$ and $0 < p < \infty$, and let $\hat{X}$ be a random variable such that

$$\mathbb{E}|\chi_{[K,\infty)}(X) - \chi_{[K,\infty)}(\hat{X})| = \varepsilon$$

for some $\varepsilon \in (0,1]$. Define $\varepsilon_1 := \mathbb{P}(X \geq K, \ \hat{X} < K)$ and $\varepsilon_2 := \mathbb{P}(X < K, \ \hat{X} \geq K)$, so that $\varepsilon = \varepsilon_1 + \varepsilon_2$. Denote by α the number $\alpha(K)$

introduced in Definition 3.3 and notice that $\alpha - \varepsilon_1 \geq 0$ and $\alpha + \varepsilon_2 \leq 1$. Now

$$
\begin{aligned}
\mathbb{E}|X - \hat{X}|^p &\geq \mathbb{E}|X - \hat{X}|^p \chi_{\{X \geq K, \hat{X} < K\} \cup \{X < K, \hat{X} \geq K\}} \\
&\geq \mathbb{E}|X - K|^p \chi_{\{X \geq K, \hat{X} < K\} \cup \{X < K, \hat{X} \geq K\}} \\
&= \mathbb{E}|X - K|^p \chi_{\{X \geq K, \hat{X} < K\}} + \mathbb{E}|X - K|^p \chi_{\{X < K, \hat{X} \geq K\}}.
\end{aligned}
$$

Since X has a bounded density, we can find a number $c_0 \in [K, \infty]$ such that $\mathbb{P}(K \leq X < c_0) = \varepsilon_1$, thus also $|\{K \leq X^* < c_0\}| = \varepsilon_1$. Note that c_0 may not be unique. But $\{K \leq X < c_0\}$ is a set of probability ε_1 where $\mathbb{E}|X - K|^p \chi_A$ is minimized over all $A \subset \{X \geq K\}$ with $\mathbb{P}(A) = \varepsilon_1$, which implies that

$$
\begin{aligned}
\mathbb{E}|X - K|^p \chi_{\{X \geq K, \hat{X} < K\}} &\geq \mathbb{E}|X - K|^p \chi_{[K, c_0)}(X) \\
&= \int_{[0,1]} |X^*(s) - K|^p \chi_{[K, c_0)}(X^*(s))\, ds = \int_{\alpha - \varepsilon_1}^{\alpha} |X^*(s) - K|^p\, ds \\
&\geq \int_0^{\varepsilon_1} |d_X(K)s|^p\, ds = \frac{d_X(K)^p \varepsilon_1^{p+1}}{p+1}
\end{aligned}
$$

and by similar arguments

$$
\mathbb{E}|X - K|^p \chi_{\{X < K, \hat{X} \geq K\}} \geq \int_{\alpha}^{\alpha + \varepsilon_2} |X^*(s) - K|^p\, ds \geq \frac{d_X(K)^p \varepsilon_2^{p+1}}{p+1}.
$$

Thus

$$
(3.2) \qquad \mathbb{E}|X - \hat{X}|^p \geq \frac{d_X(K)^p (\varepsilon_1^{p+1} + \varepsilon_2^{p+1})}{p+1} \geq \frac{d_X(K)^p \varepsilon^{p+1}}{2^p(p+1)}.
$$

Now the equation (3.2) gives

$$
\begin{aligned}
&\mathbb{E}|\chi_{[K,\infty)}(X) - \chi_{[K,\infty)}(\hat{X})| \\
&\qquad \leq 2^{\frac{p}{p+1}} (p+1)^{\frac{1}{p+1}} \left(\frac{1}{d_X(K)} \right)^{\frac{p}{p+1}} \left(\mathbb{E}|X - \hat{X}|^p \right)^{\frac{1}{p+1}}.
\end{aligned}
$$

By elementary computations we can show that

$$
2^{\frac{p}{p+1}} (p+1)^{\frac{1}{p+1}} \leq 2e^{\frac{1}{2e}} \leq 3,
$$

and keeping in mind the definition of D_X we can write

$$
\mathbb{E}|\chi_{[K,\infty)}(X) - \chi_{[K,\infty)}(\hat{X})| \leq 3 D_X(K)^{\frac{p}{p+1}} \left\| X - \hat{X} \right\|_p^{\frac{p}{p+1}}.
$$

Using the definition of X^* and the boundedness assumption for the density of X we see that $1/d_X(K) \leq \sup f_X$.

Moreover, the power $\frac{p}{p+1}$ of $\left\| X - \hat{X} \right\|_p$ is sharp. To see this, we construct an example where the lower bound given by equation (3.2) is achieved. Suppose that $\Omega = [0, 1]$ is equipped with the Lebesgue measure, $K = \frac{1}{2}$ and $\varepsilon < 1$. If we take $X(\omega) = \omega$, then X has a bounded density and $d_X(\frac{1}{2}) = 1$. Now define

$$\hat{X} = \begin{cases} X, & \text{if } \omega \in [0, \frac{1}{2} - \frac{\varepsilon}{2}) \cup (\frac{1}{2} + \frac{\varepsilon}{2}, 1], \\ X + \frac{\varepsilon}{2}, & \text{if } \omega \in [\frac{1}{2} - \frac{\varepsilon}{2}, \frac{1}{2}], \\ X - \frac{\varepsilon}{2}, & \text{if } \omega \in (\frac{1}{2}, \frac{1}{2} + \frac{\varepsilon}{2}]. \end{cases}$$

Then

$$\mathbb{E}|X - \hat{X}|^p = \mathbb{E}\left|\frac{\varepsilon}{2}\right|^p \chi_{[\frac{1}{2} - \frac{\varepsilon}{2}, \frac{1}{2} + \frac{\varepsilon}{2}]}(X) = \frac{\varepsilon^{p+1}}{2^p} = \frac{d_X(\frac{1}{2})^p \varepsilon^{p+1}}{2^p},$$

which coincides with the lower bound in equation (3.2) up to the constant. Hence the power $\frac{p}{p+1}$ of $\left\|X - \hat{X}\right\|_p$ can not be increased in the assertion (i).

Now we verify (ii). Let $\delta > 0$ and choose $\hat{X} = X - \delta$. Then

$$\begin{aligned} \mathbb{E}|\chi_{[K,\infty)}(X) - \chi_{[K,\infty)}(\hat{X})| \\ = \mathbb{P}(X \geq K, \ X - \delta < K) + \mathbb{P}(X < K, \ X - \delta \geq K) \\ = \mathbb{P}(K \leq X < K + \delta), \end{aligned}$$

so that by assumption we get, for $p > p_0$, that

$$\mathbb{P}(K \leq X < K + \delta) \leq B_X \left(\mathbb{E}|\delta|^p\right)^{\frac{1}{p+1}} \leq B_X \delta^{\frac{p}{p+1}}.$$

We let p go to infinity and conclude that

$$\mathbb{P}(K \leq X < K + \delta) \leq B_X \delta.$$

Let $N \subset \mathbb{R}$ be a null set with respect to the Lebesgue measure and let $\varepsilon > 0$. Since the Lebesgue outer measure of N is also zero, we find a sequence (I_j) of open intervals such that $N \subset \bigcup I_j$ and $\sum |I_j| \leq \varepsilon$. Let $\mathcal{L}_X$ be the law of X. Then we have

$$\mathcal{L}_X((a, b)) \leq \mathcal{L}_X([a, b]) \leq B_X |b - a|$$

and

$$\mathcal{L}_X(N) \leq \mathcal{L}_X\left(\bigcup_j I_j\right) \leq \sum_j \mathcal{L}_X(I_j) \leq B_X \sum_j |I_j| \leq B_X \varepsilon.$$

This implies that $\mathcal{L}_X(N) = 0$, so $\mathcal{L}_X$ is absolutely continuous with respect to the Lebesgue measure. By the Radon-Nikodym theorem there exists a measurable function $f : \mathbb{R} \to [0, \infty)$ such that

$$\mathcal{L}_X(M) = \int_M f(x) \, dx$$

for all measurable $M \subseteq \mathbb{R}$. Moreover, f is integrable since $\mathcal{L}_X(\mathbb{R}) = 1$. Define a function $\Phi : \mathbb{R} \to [0, 1]$ such that

$$\Phi(t) = \int_{-\infty}^t f(x) \, dx.$$

Then by [18, Thm. 8.17] we have that $\Phi'(t) = f(t)$ a.e. in $\mathbb{R}$. On the other hand, we have that

$$\Phi'(t) = \lim_{h \to 0} \frac{\Phi(t+h) - \Phi(t)}{h} \leq \lim_{h \to 0} \frac{B_X h}{h} = B_X \text{ a.e. in } \mathbb{R},$$

because $\Phi(t+h) - \Phi(t) = \mathcal{L}_X((t, t+h))$. Therefore we conclude that $f(t) \leq B_X$ a.e. in $\mathbb{R}$. $\qquad\square$

Remark 3.5. By considering complements of the intervals in the indicator functions and the random variables $-X$ and $-\hat{X}$, we have corresponding results for the functions $\chi_{(K,\infty)}$, $\chi_{(-\infty,K]}$ and $\chi_{(-\infty,K)}$.

As an immediate consequence of Theorem 3.4, we can derive

Corollary 3.6. *Let X be the solution of the equation* (2.1), *$K \in \mathbb{R}$ and $0 < p < \infty$. Let X_T have a bounded density and suppose that X_T^π is an approximation of X_T such that*

$$\|X_T - X_T^\pi\|_p \leq C_p \, |\pi|^\theta$$

for some $\theta > 0$ and some constant $C_p \geq 0$. Then for all $K \in \mathbb{R}$ we have

$$\mathbb{E}|\chi_{[K,\infty)}(X_T) - \chi_{[K,\infty)}(X_T^\pi)| \leq 3 D_{X_T}(K)^{\frac{p}{p+1}} C_p^{\frac{p}{p+1}} \, |\pi|^{\frac{\theta p}{p+1}} .$$

3.2. **Euler and Milstein Schemes.** Now we can apply the results of Section 3.1 to the Euler and Milstein schemes:

Theorem 3.7. *For any $0 < \varepsilon < 1/2$ there exists a constant $C_\varepsilon > 0$ such that for all $K \in \mathbb{R}$ we have that*

$$\mathbb{E}|\chi_{[K,\infty)}(X_T) - \chi_{[K,\infty)}(X_T^E)| \leq (D_{X_T}(K) \vee \sqrt{D_{X_T}(K)}) C_\varepsilon \, |\pi|^{\frac{1}{2} - \varepsilon}$$

and for any $0 < \varepsilon < 1$ there exists a constant $C'_\varepsilon > 0$ such that for all $K \in \mathbb{R}$ we have that

$$\mathbb{E}|\chi_{[K,\infty)}(X_T) - \chi_{[K,\infty)}(X_T^M)| \leq (D_{X_T}(K) \vee \sqrt{D_{X_T}(K)}) C'_\varepsilon \, |\pi|^{1-\varepsilon} .$$

Proof. Let $1 \leq p < \infty$. Then for the Euler scheme we have by Theorem A.1 in the Appendix that

$$\left\|X_T - X_T^E\right\|_p \leq C_p \, |\pi|^{\frac{1}{2}} ,$$

i.e. the assumption of Corollary 3.6 is satisfied with $\theta = \frac{1}{2}$. Thus

$$(3.3) \qquad \mathbb{E}|\chi_{[K,\infty)}(X_T) - \chi_{[K,\infty)}(X_T^E)| \leq 3 D_{X_T}(K)^{\frac{p}{p+1}} C_p^{\frac{p}{p+1}} \, |\pi|^{\frac{p}{2(p+1)}} .$$

Similarly for the Milstein scheme we have by [13, Proposition 1, p. 140] that

$$\left\|X_T - X_T^M\right\|_p \leq C'_p \, |\pi| ,$$

which gives the assumption of Corollary 3.6 with $\theta = 1$, and therefore

$$(3.4) \qquad \mathbb{E}|\chi_{[K,\infty)}(X_T) - \chi_{[K,\infty)}(X_T^M)| \leq 3 D_{X_T}(K)^{\frac{p}{p+1}} (C'_p)^{\frac{p}{p+1}} \, |\pi|^{\frac{p}{p+1}} .$$

The claim follows in both cases by choosing p such that $p = (\theta - \varepsilon)/\varepsilon$, where $0 < \varepsilon < \theta$, and noticing that for any $a > 0$ we have $a^{\frac{p}{p+1}} \leq a \vee \sqrt{a}$. The constant 3 and the constants coming from the approximation schemes are included in C_ε or C'_ε, which now depend on ε through the choice of p. $\qquad\square$

Since we have information about the constant C_p in Theorem A.1, i.e. $C_p = e^{Mp^2}$, we can write an extended version of Theorem 3.7 for the Euler scheme:

Theorem 3.8. *Let $K \in \mathbb{R}$. Then there exists $m \in (0,1)$ such that for $|\pi| < m$ we have*

$$\mathbb{E}|\chi_{[K,\infty)}(X_T) - \chi_{[K,\infty)}(X_T^E)| \leq (D_{X_T}(K) \vee \sqrt{D_{X_T}(K)}) |\pi|^{\frac{1}{2} - \frac{2+M}{(-\log|\pi|)^{1/3}}},$$

where the constant $M = M(x_0, T, C_T) \in (0, \infty)$ is taken from Theorem A.1.

Proof. By Corollary 3.6 and Theorem A.1, using $a^{\frac{p}{p+1}} \leq a \vee \sqrt{a}$ for $a > 0$ and $p \geq 1$, we get

$$\mathbb{E}|\chi_{[K,\infty)}(X_T) - \chi_{[K,\infty)}(X_T^E)| \leq 3 D_{X_T}(K)^{\frac{p}{p+1}} C_p^{\frac{p}{p+1}} |\pi|^{\frac{p}{2(p+1)}}$$

(3.5)
$$\leq (D_{X_T}(K) \vee \sqrt{D_{X_T}(K)}) 3 e^{Mp^2 \cdot \frac{p}{p+1}} |\pi|^{\frac{p}{2(p+1)}}$$

$$\leq (D_{X_T}(K) \vee \sqrt{D_{X_T}(K)}) 3 e^{Mp^2} |\pi|^{\frac{p}{2(p+1)}}$$

for all $p \geq 1$. Now choose p such that

$$4p(p+1)^2 = -\log|\pi|$$

for $|\pi| \leq m$ with $m = e^{-16}$. This gives $p^3 \leq -\log|\pi|$ and $p^2 \leq (-\log|\pi|)^{2/3}$. Thus we have

$$e^{Mp^2} \leq e^{M(-\log|\pi|)^{2/3}} = |\pi|^{-M(-\log|\pi|)^{-1/3}}$$

and

$$\frac{1}{2(p+1)} = \sqrt{\frac{p}{-\log|\pi|}} \leq \sqrt{(-\log|\pi|)^{1/3-1}} = (-\log|\pi|)^{-1/3}.$$

Using these we get

$$3 e^{Mp^2} |\pi|^{\frac{p}{2(p+1)}} = 3 e^{Mp^2} |\pi|^{\frac{1}{2} - \frac{1}{2(p+1)}} \leq 3 |\pi|^{\frac{1}{2} - \frac{1+M}{(-\log|\pi|)^{1/3}}} \leq |\pi|^{\frac{1}{2} - \frac{2+M}{(-\log|\pi|)^{1/3}}},$$

where in the last step we used the inequality

$$3 |\pi|^{\frac{1}{(-\log|\pi|)^{1/3}}} \leq 1$$

for $|\pi| < m$. Now we come back to equation (3.5) and conclude that

$$\mathbb{E}|\chi_{[K,\infty)}(X_T) - \chi_{[K,\infty)}(X_T^E)|$$

$$\leq (D_{X_T}(K) \vee \sqrt{D_{X_T}(K)}) |\pi|^{\frac{1}{2} - \frac{2+M}{(-\log|\pi|)^{1/3}}}.$$

$\qquad\square$

4. Functions of Bounded Variation

From Theorem 3.4 we deduce the same error for functions of bounded variation, up to a constant. Let us first recall the definitions of the spaces BV and NBV.

Definition 4.1. Let

$$T_f(x) := \sup \sum_{j=1}^{N} |f(x_j) - f(x_{j-1})|,$$

where the supremum is taken over N and all partitions $-\infty < x_0 < x_1 < \ldots < x_N = x$, be the total variation function of f. Then we say that f is a function of bounded variation, $f \in BV$, if

$$V(f) := \lim_{x \to \infty} T_f(x)$$

is finite, and call $V(f)$ the (total) variation of f.

Definition 4.2. Let NBV be the set of functions $f \in BV$ such that f is left-continuous and $f(x) \to 0$ as $x \to -\infty$.

4.1. General Approximation.

Theorem 4.3. *Suppose that X and $\hat{X}$ are random variables and X has a bounded density. If $g \in BV$ and $1 \leq p < \infty$, then for any $1 \leq q < \infty$ we have*

$$\mathbb{E}|g(X) - g(\hat{X})|^p \leq 3^{p+1} \left(\sup f_X\right)^{\frac{q}{q+1}} V(g)^p \left\|X - \hat{X}\right\|_q^{\frac{q}{q+1}}.$$

Proof. First we show the result for functions $g \in NBV$. By [18, Thm. 8.14] there is a unique signed measure μ such that

$$g(x) = \mu((-\infty, x)) \text{ and } |\mu|((-\infty, x)) = T_g(x),$$

where $|\mu|$ is the total variation measure of μ. We consider the Jordan decomposition of μ, i.e. $\mu = \mu_1 - \mu_2$, where $\mu_1 = \frac{1}{2}(|\mu| + \mu)$ and $\mu_2 = \frac{1}{2}(|\mu| - \mu)$ are positive measures. Then $|\mu| = \mu_1 + \mu_2$, and all three measures $|\mu|$, μ_1 and μ_2 are finite since $|\mu|(\mathbb{R}) = V(g) < \infty$. Thus we get

$$g(x) = \mu((-\infty, x)) = \int_{\mathbb{R}} \chi_{(-\infty, x)}(z) \, d\mu(z) = \int_{\mathbb{R}} \chi_{(z, \infty)}(x) \, d\mu(z).$$

Now by Theorem 3.4 and Remark 3.5 we get

$$\left\|g(X) - g(\hat{X})\right\|_p = \left\|\int_{\mathbb{R}} \chi_{(z,\infty)}(X) \, d\mu(z) - \int_{\mathbb{R}} \chi_{(z,\infty)}(\hat{X}) \, d\mu(z)\right\|_p$$

$$= \left\|\int_{\mathbb{R}} \left[\chi_{(z,\infty)}(X) - \chi_{(z,\infty)}(\hat{X})\right] \, d\mu(z)\right\|_p$$

$$\leq \left\| \int_{\mathbb{R}} \left| \chi_{(z,\infty)}(X) - \chi_{(z,\infty)}(\hat{X}) \right| d|\mu|(z) \right\|_p$$

$$\leq \int_{\mathbb{R}} \left\| \chi_{(z,\infty)}(X) - \chi_{(z,\infty)}(\hat{X}) \right\|_p d|\mu|(z)$$

$$\leq 3^{\frac{1}{p}} (\sup f_X)^{\frac{q}{p(q+1)}} V(g) \left\| X - \hat{X} \right\|_q^{\frac{q}{p(q+1)}},$$

which completes the proof for functions in NBV.

Next, let g be an arbitrary function in BV. By [18, Thm. 8.13], there exists a unique function $\tilde{g} \in NBV$ and a unique constant $c \in \mathbb{R}$ such that $g(x) = \tilde{g}(x) + c$ at all points of continuity of g, with $V(\tilde{g}) \leq V(g)$. Also by [18] we know that g can have only countably many points of discontinuity, so define $\cup_{j=1}^{\infty}\{a_j\}$ to be the set of these points and let $\lambda_j := g(a_j) - \tilde{g}(a_j) - c$. Then we can write

$$g(x) = \tilde{g}(x) + c + \Delta(x),$$

where

$$\Delta(x) := \sum_{j=1}^{\infty} \lambda_j \chi_{\{a_j\}}(x) = \sum_{j=1}^{\infty} \lambda_j \left(\chi_{(-\infty,a_j]}(x) - \chi_{(-\infty,a_j)}(x) \right).$$

We define a measure

$$\nu = \sum_{j=1}^{\infty} \lambda_j \delta_{a_j},$$

where δ_a is the Dirac measure in a. Again by [18] we know that $g(a_j-)$ exists, so we have $\tilde{g}(a_j) + c = g(a_j-)$ and

$$|\nu|(\mathbb{R}) = \sum_{j=1}^{\infty} |\lambda_j| = \sum_{j=1}^{\infty} |g(a_j) - g(a_j-)| \leq V(g).$$

Now we can write

$$\Delta(x) = \int_{\mathbb{R}} \chi_{(-\infty,z]}(x) - \chi_{(-\infty,z)}(x) \, d\nu(z).$$

and compute, similarly as in the NBV case, that

$$\left\| \Delta(X) - \Delta(\hat{X}) \right\|_p \leq \left\| \int_{\mathbb{R}} |\chi_{(-\infty,z]}(X) - \chi_{(-\infty,z]}(\hat{X})| \, d|\nu|(z) \right\|_p$$

$$+ \left\| \int_{\mathbb{R}} |\chi_{(-\infty,z)}(X) - \chi_{(-\infty,z)}(\hat{X})| \, d|\nu|(z) \right\|_p$$

$$\leq \int_{\mathbb{R}} \left\| \chi_{(-\infty,z]}(X) - \chi_{(-\infty,z]}(\hat{X}) \right\|_p d|\nu|(z)$$

$$+ \int_{\mathbb{R}} \left\| \chi_{(-\infty,z)}(X) - \chi_{(-\infty,z)}(\hat{X}) \right\|_p d|\nu|(z)$$

$$\leq 2 \cdot 3^{\frac{1}{p}} (\sup f_X)^{\frac{q}{p(q+1)}} V(g) \left\| X - \hat{X} \right\|_q^{\frac{q}{p(q+1)}}.$$

This, combined with the NBV result, implies that

$$\left\|g(X) - g(\hat{X})\right\|_p = \left\|\tilde{g}(X) - \tilde{g}(\hat{X}) + \Delta(X) - \Delta(\hat{X})\right\|_p$$

$$\leq \left\|\tilde{g}(X) - \tilde{g}(\hat{X})\right\|_p + \left\|\Delta(X) - \Delta(\hat{X})\right\|_p$$

$$\leq 3 \cdot 3^{\frac{1}{p}} (\sup f_X)^{\frac{q}{p(q+1)}} V(g) \left\|X - \hat{X}\right\|_q^{\frac{q}{p(q+1)}},$$

which gives the statement. □

As in Corollary 3.6 for indicator functions, we can now write an analoguous statement for functions of bounded variation:

Corollary 4.4. *Let X be the solution of the equation (2.1), $1 \leq p < \infty$ and $g \in BV$. Suppose that X_T has a bounded density, $1 \leq q < \infty$ and X_T^π is an approximation of X_T such that*

$$\|X_T - X_T^\pi\|_q \leq C_q |\pi|^\theta$$

for some $\theta > 0$ and some constant $C_q \geq 0$. Then

$$\mathbb{E}|g(X_T) - g(X_T^\pi)|^p \leq 3^{p+1} (\sup f_{X_T})^{\frac{q}{q+1}} V(g)^p C_q^{\frac{q}{q+1}} |\pi|^{\frac{\theta q}{q+1}}.$$

4.2. **Euler and Milstein Schemes.**

Theorem 4.5. *Let $g \in BV$ and $1 \leq p < \infty$. Then we have for $0 < \varepsilon < 1/2$ that*

$$\mathbb{E}|g(X_T) - g(X_T^E)|^p \leq 3^p (\sup f_{X_T} \vee \sqrt{\sup f_{X_T}}) V(g)^p C_\varepsilon |\pi|^{\frac{1}{2}-\varepsilon}$$

and for $0 < \varepsilon < 1$ that

$$\mathbb{E}|g(X_T) - g(X_T^M)|^p \leq 3^p (\sup f_{X_T} \vee \sqrt{\sup f_{X_T}}) V(g)^p C_\varepsilon' |\pi|^{1-\varepsilon},$$

where C_ε and C_ε' depend on ε and the constants of the corresponding schemes.

Proof. The statement follows from Corollary 4.4 with arguments similar to the proof of Theorem 3.7. □

For the Euler scheme we can write an extended version corresponding to Theorem 3.8:

Theorem 4.6. *Let $1 \leq p < \infty$ and $g \in BV$. Then there exists $m > 0$ such that for $|\pi| < m$ we have*

$$\mathbb{E}|g(X_T) - g(X_T^E)|^p \leq 3^p (\sup f_{X_T} \vee \sqrt{\sup f_{X_T}}) V(g)^p |\pi|^{\frac{1}{2} - \frac{2+M}{(-\log|\pi|)^{1/3}}},$$

where M is the constant in Theorem A.1.

Proof. By Theorem A.1 and Corollary 4.4 we get for $1 \leq q < \infty$ that

$$\mathbb{E}|g(X_T) - g(X_T^E)|^p \leq 3^{p+1} (\sup f_{X_T} \vee \sqrt{\sup f_{X_T}}) V(g)^p e^{Mq^2 \cdot \frac{q}{q+1}} |\pi|^{\frac{q}{2(q+1)}}$$

$$\leq 3^{p+1} (\sup f_{X_T} \vee \sqrt{\sup f_{X_T}}) V(g)^p e^{Mq^2} |\pi|^{\frac{q}{2(q+1)}},$$

which by the arguments in Theorem 3.8 implies that

$$\mathbb{E}|g(X_T) - g(X_T^E)|^p \leq 3^p(\sup f_{X_T} \vee \sqrt{\sup f_{X_T}})V(g)^p |\pi|^{\frac{1}{2} - \frac{2+M}{(-\log|\pi|)^{1/3}}}.$$

□

5. Extension

Next we extend the result into a function class, to be called $\mathcal{G}_{p,\varphi}$, that contains e.g. all polynomials. The main result is given in Theorem 5.7 and the class $\mathcal{G}_{p,\varphi}$ is analyzed in Section 6.

Definition 5.1 (Bump function). Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a function such that $0 < \varphi(z) \leq 1$ for all $z \in \mathbb{R}$, φ is increasing in $(-\infty, 0]$ and decreasing in $(0, \infty)$, and

$$\lim_{|z| \to \infty} \varphi(z) = 0.$$

Then φ is called a bump function.

Definition 5.2 (Class $\mathcal{G}_{p,\varphi}$). Fix $p \in [1, \infty)$ and let $\varphi : \mathbb{R} \to \mathbb{R}$ be a given bump function. Let $\mathcal{M}$ be the set of all signed measures μ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that $|\mu|$ is σ-finite. Define

$$\mathcal{M}_{p,\varphi} = \{\mu \in \mathcal{M} : \varphi \in L_1^{loc}(\mathbb{R}, |\mu|) \cap L_{1+\frac{1}{p}}(\mathbb{R}, |\mu|)\}.$$

Then for any $\mu \in \mathcal{M}_{p,\varphi}$ define a function related to μ by

$$g^\mu(x) = \begin{cases} \int_{(0,x]} \varphi \, d\mu, & \text{for } x \geq 0, \\ \int_{(x,0]} \varphi \, d\mu, & \text{for } x < 0, \end{cases}$$

where $(0, 0] = \emptyset$. Also define a set of jump functions

$$\Delta_{p,\varphi} = \{ \ \Delta_A(x) = \sum_{i=1}^{\infty} \lambda_i \varphi(a_i) \chi_{\{a_i\}}(x) : A = (a_i)_{i=1}^{\infty} \subset \mathbb{R} \text{ countable },$$

$$a_i \neq a_j \text{ if } i \neq j, \ (\lambda_i)_{i=1}^{\infty} \subset \mathbb{R} \text{ and } \sum_{i=1}^{\infty} |\lambda_i| \varphi(a_i)^{1+\frac{1}{p}} < \infty\}.$$

Then we set

$$\mathcal{G}_{p,\varphi} = \{g : \ g = c + g^\mu + \Delta_A \text{ for some } c \in \mathbb{R},$$
$$\mu \in \mathcal{M}_{p,\varphi} \text{ and } \Delta_A \in \Delta_{p,\varphi}\},$$

where the decomposition of g is unique, as we will see in Theorem 5.4. Moreover, denote the (p, φ)-variation of $g \in \mathcal{G}_{p,\varphi}$ by

$$V_{p,\varphi}(g) = \int_{\mathbb{R}} \varphi^{1+\frac{1}{p}} \, d|\mu| + \sum_{i=1}^{\infty} |\lambda_i| \varphi(a_i)^{1+\frac{1}{p}}.$$

Remark 5.3. The definition implies that any function $g^\mu \in \mathcal{G}_{p,\varphi}$ is right-continuous and $g^\mu(0) = 0$. To relax these restrictions, we add to the function g^μ a constant c and a function Δ_A, which can be used to alter the left- or right-continuity of g^μ at the points of discontinuity or to add point discontinuities anywhere. For example, we can make g^μ left-continuous by choosing

$$\Delta_A(x) = \begin{cases} g^\mu(x-) - g^\mu(x) \text{ for } x \in A, \\ 0 \text{ elsewhere,} \end{cases}$$

where A is the set of points of discontinuity of g^μ. We see that g can have only a countable number of jumps, because otherwise $\mu(\{x\}) \neq 0$ for uncountably many $x \in \mathbb{R}$, which is a contradiction to the σ-finiteness of the measure $|\mu|$. Moreover, both g and g^μ may jump at zero; the jump height of g^μ is then given by $\varphi(0)\mu(\{0\})$, and the jump of g depends on $\varphi(0)\mu(\{0\})$ and $\Delta_A(0)$.

Theorem 5.4. *The decomposition $g = c + g^\mu + \Delta_A$ for functions $g \in \mathcal{G}_{p,\varphi}$ is unique.*

Proof. Take $g_1, g_2 \in \mathcal{G}_{p,\varphi}$ such that $g_i = c_i + g^{\mu_i} + \Delta_{A_i}$, $i \in \{1,2\}$, and suppose that $g_1 = g_2$. Now $A_1 \cup A_2$ is countable and $\Delta_{A_i} = 0$ in $(A_1 \cup A_2)^c$. Let us take a sequence $(x_j) \subset (A_1 \cup A_2)^c$ such that $x_j \searrow 0$ as $j \to \infty$. Since g^{μ_i} is right-continuous and $g^{\mu_i}(0) = 0$, we get that $g_i(x_j) = c_i + g^{\mu_i}(x_j) \to c_i$, and thus $c_1 = c_2$. This implies that for $x_0 \in (A_1 \cup A_2)^c$ we have $g^{\mu_1}(x_0) = g^{\mu_2}(x_0)$. Now let $x_0 \in A_1 \cup A_2$. Again we choose a sequence $(x_j) \subset (A_1 \cup A_2)^c$ such that $x_j \searrow x_0$ as $j \to \infty$, and by right-continuity of g^{μ_i} we get that $g^{\mu_1}(x_0) = g^{\mu_2}(x_0)$. Thus $g^{\mu_1} = g^{\mu_2}$ everywhere, and also $\Delta_{A_1} = \Delta_{A_2}$. $\square$

Theorem 5.5. *Functions of bounded variation are a special case of functions in $\mathcal{G}_{p,\varphi}$, i.e. $BV \subset \mathcal{G}_{p,\varphi}$.*

Proof. Let $g \in NBV$ and let μ^{BV} be the signed measure related to g. At the points of continuity of g we have $g = g(0+) + g^\mu$, where the measure μ is chosen such that $d\mu = d\mu^{BV}/\varphi$ on $(0, \infty)$ and $d\mu = -d\mu^{BV}/\varphi$ on $(-\infty, 0]$. Here $|\mu|$ is σ-finite by the finiteness of μ^{BV} and the properties of φ. It also holds that

$$\|\varphi\|_{L_{1+\frac{1}{p}}(\mathbb{R}, |\mu|)} \leq V(g)^{\frac{p}{p+1}} < \infty.$$

Now let $g \in BV$. Then $g = \tilde{g} + c$ for some $\tilde{g} \in NBV$ and $c \in \mathbb{R}$ at the points of continuity of g, thus satisfying $g = \tilde{g}(0+) + c + \tilde{g}^\mu$. At the points of discontinuity we correct this by choosing Δ_A such that A is the set of the points where g is not right-continuous, and the values λ correspond to the jump heights of g. Then

$$\sum_{i=1}^\infty |\lambda_i| \varphi(a_i)^{1+\frac{1}{p}} \leq \sum_{i=1}^\infty |\lambda_i| \leq V(g).$$

$\square$

5.1. **General Approximation.** As before, let X and $\hat{X}$ be random variables. We define a function φ that connects the random variables with their tail behavior.

Definition 5.6. Take two strictly positive monotone functions

$$\varphi_\theta^+ : (0, \infty) \to (0, 1]$$

and

$$\varphi_\theta^- : (-\infty, 0] \to (0, 1]$$

for $0 < \theta < 1$ with properties

$$\varphi_\theta^+(K) \to 0 \text{ as } K \to \infty,$$

$$\varphi_\theta^-(K) \to 0 \text{ as } K \to -\infty,$$

$$\left[\mathbb{P}(X \geq K) \vee \mathbb{P}(\hat{X} \geq K)\right]^\theta \leq \varphi_\theta^+(K) \text{ for } K > 0$$

and

$$\left[\mathbb{P}(X \leq K) \vee \mathbb{P}(\hat{X} \leq K)\right]^\theta \leq \varphi_\theta^-(K) \text{ for } K \leq 0.$$

Then we define a bump function $\varphi_\theta^{X,\hat{X}} : \mathbb{R} \to (0, 1]$ by

$$\varphi_\theta^{X,\hat{X}}(K) := \begin{cases} \varphi_\theta^+(K) & \text{if } K > 0, \\ \varphi_\theta^-(K) & \text{if } K \leq 0. \end{cases}$$

Now the main result is the following convergence theorem for functions in the class $\mathcal{G}_{p,\varphi}$ associated with the function $\varphi_\theta^{X,\hat{X}}$.

Theorem 5.7. *Let X and $\hat{X}$ be random variables such that X has a bounded density. Suppose that $0 < \theta < 1$ and let $\varphi_\theta^{X,\hat{X}}$ be a function as in Definition 5.6. If $1 \leq p < \infty$ and $g \in \mathcal{G}_{p,\varphi_\theta^{X,\hat{X}}}$, then for all $1 \leq q < \infty$ we have*

$$\mathbb{E}|g(X) - g(\hat{X})|^p \leq 3 \cdot 2^p \left(\sup f_X\right)^{\frac{q(1-\theta)}{q+1}} \left(V_{p,\varphi_\theta^{X,\hat{X}}}(g)\right)^p \left\|X - \hat{X}\right\|_q^{\frac{q(1-\theta)}{q+1}}.$$

Proof. Let $g \in \mathcal{G}_{p,\varphi_\theta^{X,\hat{X}}}$. Then by definition $g = c + g^\mu + \Delta_A$ and

$$\left\|g(X) - g(\hat{X})\right\|_p \leq \left\|g^\mu(X) - g^\mu(\hat{X})\right\|_p + \left\|\Delta_A(X) - \Delta_A(\hat{X})\right\|_p.$$

Now we can compute

$$
\begin{aligned}
g^\mu(x)\chi_{[0,\infty)}(x) &= \int_{(0,x]} \varphi_\theta^{X,\hat{X}}(z)\, d\mu(z)\, \chi_{[0,\infty)}(x) \\
&= \int_{(0,\infty)} \chi_{[z,\infty)}(x)\varphi_\theta^{X,\hat{X}}(z)\, d\mu(z)
\end{aligned}
$$

and similarly

$$g^\mu(x)\chi_{(-\infty,0)}(x) = \int_{(x,0]} \varphi_\theta^{X,\hat{X}}(z)\,d\mu(z)\,\chi_{(-\infty,0)}(x)$$

$$= \int_{(-\infty,0]} \chi_{(-\infty,z)}(x)\varphi_\theta^{X,\hat{X}}(z)\,d\mu(z).$$

Thus

$$\left\|g^\mu(X) - g^\mu(\hat{X})\right\|_p \le \left\|g^\mu(X)\chi_{[0,\infty)}(X) - g^\mu(\hat{X})\chi_{[0,\infty)}(\hat{X})\right\|_p$$

$$+ \left\|g^\mu(X)\chi_{(-\infty,0)}(X) - g^\mu(\hat{X})\chi_{(-\infty,0)}(\hat{X})\right\|_p$$

$$\le \left\|\int_{(0,\infty)} |\chi_{[z,\infty)}(X) - \chi_{[z,\infty)}(\hat{X})|\varphi_\theta^{X,\hat{X}}(z)\,d|\mu|(z)\right\|_p$$

$$+ \left\|\int_{(-\infty,0]} |\chi_{(-\infty,z)}(X) - \chi_{(-\infty,z)}(\hat{X})|\varphi_\theta^{X,\hat{X}}(z)\,d|\mu|(z)\right\|_p$$

$$\le \int_{(0,\infty)} \left\|\chi_{[z,\infty)}(X) - \chi_{[z,\infty)}(\hat{X})\right\|_p \varphi_\theta^{X,\hat{X}}(z)\,d|\mu|(z)$$

$$+ \int_{(-\infty,0]} \left\|\chi_{(-\infty,z)}(X) - \chi_{(-\infty,z)}(\hat{X})\right\|_p \varphi_\theta^{X,\hat{X}}(z)\,d|\mu|(z).$$

Denote by $\psi(X,\hat{X})$ the error function from Theorem 3.4, i.e.

$$\psi(X,\hat{X}) := 3(\sup f_X)^{\frac{q}{q+1}} \left\|X - \hat{X}\right\|_q^{\frac{q}{q+1}}$$

and notice that $a \wedge b \le a^{1-\theta}b^\theta$ for any $a,b \ge 0$ and $0 < \theta < 1$. Since

$$\mathbb{E}|\chi_{[K,\infty)}(X) - \chi_{[K,\infty)}(\hat{X})| = \mathbb{P}(X \ge K, \hat{X} < K) + \mathbb{P}(X < K, \hat{X} \ge K)$$

$$\le 2(\mathbb{P}(X \ge K) \vee \mathbb{P}(\hat{X} \ge K)),$$

it follows from Theorem 3.4 that, for $K > 0$,

$$\mathbb{E}|\chi_{[K,\infty)}(X) - \chi_{[K,\infty)}(\hat{X})| \le \psi(X,\hat{X}) \wedge 2\left[\mathbb{P}(X \ge K) \vee \mathbb{P}(\hat{X} \ge K)\right]$$

$$\le \psi(X,\hat{X})^{1-\theta}2^\theta\left[\mathbb{P}(X \ge K) \vee \mathbb{P}(\hat{X} \ge K)\right]^\theta$$

$$\le 2^\theta\psi(X,\hat{X})^{1-\theta}\varphi_\theta^+(K),$$

where $0 < \theta < 1$. In a similar way we get for $K \le 0$ that

$$\mathbb{E}|\chi_{[K,\infty)}(X) - \chi_{[K,\infty)}(\hat{X})| \le 2^\theta\psi(X,\hat{X})^{1-\theta}\varphi_\theta^-(K),$$

so we can write for $K \in \mathbb{R}$ that

$$\mathbb{E}|\chi_{[K,\infty)}(X) - \chi_{[K,\infty)}(\hat{X})| \le 2^\theta\psi(X,\hat{X})^{1-\theta}\varphi_\theta^{X,\hat{X}}(K).$$

This gives an estimate for $\left\|\chi_{[z,\infty)}(X) - \chi_{[z,\infty)}(\hat{X})\right\|_p$, and the same estimate holds for $\left\|\chi_{(-\infty,z)}(X) - \chi_{(-\infty,z)}(\hat{X})\right\|_p$ by the observation in Remark 3.5. Therefore

$$\left\|g^\mu(X) - g^\mu(\hat{X})\right\|_p \leq \int_{\mathbb{R}} 2^{\frac{\theta}{p}} \psi(X,\hat{X})^{\frac{1-\theta}{p}} \varphi_\theta^{X,\hat{X}}(z)^{\frac{1}{p}} \varphi_\theta^{X,\hat{X}}(z)\, d|\mu|(z)$$

$$\leq 2^{\frac{\theta}{p}} \int_{\mathbb{R}} 3^{\frac{1-\theta}{p}} (\sup f_X)^{\frac{q(1-\theta)}{(q+1)p}} \left\|X - \hat{X}\right\|_q^{\frac{q(1-\theta)}{(q+1)p}} \varphi_\theta^{X,\hat{X}}(z)^{1+\frac{1}{p}}\, d|\mu|(z)$$

$$\leq 3^{\frac{1}{p}} (\sup f_X)^{\frac{q(1-\theta)}{(q+1)p}} \int_{\mathbb{R}} \varphi_\theta^{X,\hat{X}}(z)^{1+\frac{1}{p}}\, d|\mu|(z) \left\|X - \hat{X}\right\|_q^{\frac{q(1-\theta)}{(q+1)p}}.$$

It remains to show a similar estimate for the jump function Δ_A. This can be done by the same argument as in the case of bounded variation, namely by writing

$$\Delta_A(x) = \int_{\mathbb{R}} \chi_{(-\infty,z]}(x) - \chi_{(-\infty,z)}(x)\, d\nu(z),$$

where

$$\nu = \sum_{i=1}^{\infty} \lambda_i \varphi_\theta^{X,\hat{X}}(a_i) \delta_{a_i}$$

and δ_a is the Dirac measure in a. Then by arguments similar to the first part of the proof and Remark 3.5 we get

$$\left\|\Delta_A(X) - \Delta_A(\hat{X})\right\|_p \leq \left\|\int_{\mathbb{R}} |\chi_{(-\infty,z]}(X) - \chi_{(-\infty,z]}(\hat{X})|\, d|\nu|(z)\right\|_p$$

$$+ \left\|\int_{\mathbb{R}} |\chi_{(-\infty,z)}(X) - \chi_{(-\infty,z)}(\hat{X})|\, d|\nu|(z)\right\|_p$$

$$\leq \int_{\mathbb{R}} \left\|\chi_{(-\infty,z]}(X) - \chi_{(-\infty,z]}(\hat{X})\right\|_p d|\nu|(z)$$

$$+ \int_{\mathbb{R}} \left\|\chi_{(-\infty,z)}(X) - \chi_{(-\infty,z)}(\hat{X})\right\|_p d|\nu|(z)$$

$$\leq 2 \cdot 3^{\frac{1}{p}} (\sup f_X)^{\frac{q(1-\theta)}{(q+1)p}} \int_{\mathbb{R}} \varphi_\theta^{X,\hat{X}}(z)^{\frac{1}{p}}\, d|\nu|(z) \left\|X - \hat{X}\right\|_q^{\frac{q(1-\theta)}{(q+1)p}}$$

$$\leq 2 \cdot 3^{\frac{1}{p}} (\sup f_X)^{\frac{q(1-\theta)}{(q+1)p}} \left(\sum_{i=1}^{\infty} |\lambda_i| \varphi_\theta^{X,\hat{X}}(a_i)^{1+\frac{1}{p}}\right) \left\|X - \hat{X}\right\|_q^{\frac{q(1-\theta)}{(q+1)p}},$$

so finally we get

$$\left\|g(X) - g(\hat{X})\right\|_p \leq 2 \cdot 3^{\frac{1}{p}} (\sup f_X)^{\frac{q(1-\theta)}{(q+1)p}} V_{p,\varphi_\theta^{X,\hat{X}}}(g) \left\|X - \hat{X}\right\|_q^{\frac{q(1-\theta)}{(q+1)p}}.$$

$\square$

6. Analysis of the Class $\mathcal{G}_{p,\varphi}$

We study the class $\mathcal{G}_{p,\varphi}$ with the underlying function $\varphi = \varphi_\theta^{X,\hat{X}}$. This function depends on the approximation $\hat{X}$, and our first task is to handle this dependence. We show in Lemma 6.2 that we can choose the function $\varphi_\theta^{X,\hat{X}}$ such that it decays faster than any polynomial, and then we prove in Theorem 6.3 that with this choice, the class $\mathcal{G}_{p,\varphi_\theta^{X,\hat{X}}}$ contains all polynomials. Then we apply the results to solutions of SDEs, and collect our knowledge in the main result, Corollary 6.4.

Lemma 6.1. *Suppose that φ and ψ are bump functions.*

(i) *If $\varphi \le \psi$, then $\mathcal{G}_{p,\psi} \subset \mathcal{G}_{p,\varphi}$.*
(ii) *If $g \in \mathcal{G}_{p,\psi}$ and $\varphi \le \psi$, then $V_{p,\varphi}(g) \le V_{p,\psi}(g)$.*

Proof. First we show (i). Let $g \in \mathcal{G}_{p,\psi}$ and μ_ψ be related to g, i.e. $g = c + g^{\mu_\psi} + \Delta_A$. We choose a measure μ_φ such that $d\mu_\varphi = (\psi/\varphi)d\mu_\psi$, which implies that $d|\mu_\varphi| = (\psi/\varphi)d|\mu_\psi|$ and $|\mu_\varphi|$ is σ-finite. Now we get for $x \ge 0$ that

$$g^{\mu_\psi}(x) = \int_{(0,x]} \psi(z)\, d\mu_\psi(z) = \int_{(0,x]} \varphi(z)\, d\mu_\varphi(z)$$

and similarly for $x < 0$. The integrability conditions are satisfied, since

$$\int_A \varphi(z)\, d|\mu_\varphi|(z) = \int_A \psi(z)\, d|\mu_\psi|(z) < \infty$$

for all $A \subset\subset \mathbb{R}$ and

$$\int_{\mathbb{R}} \varphi(z)^{1+\frac{1}{p}}\, d|\mu_\varphi|(z) = \int_{\mathbb{R}} \psi(z)\varphi(z)^{\frac{1}{p}}\, d|\mu_\psi|(z)$$
$$\le \int_{\mathbb{R}} \psi(z)^{1+\frac{1}{p}}\, d|\mu_\psi|(z) < \infty.$$

The representation of the jump part Δ_A changes correspondingly in the change of measure, i.e. we set

$$\lambda_i^\varphi = \lambda_i^\psi \cdot \frac{\psi(a_i)}{\varphi(a_i)}$$

and see that

$$\sum_{i=1}^\infty |\lambda_i^\varphi|\varphi(a_i)^{1+\frac{1}{p}} \le \sum_{i=1}^\infty |\lambda_i^\psi|\psi(a_i)^{1+\frac{1}{p}} < \infty.$$

This proves the assertion (i), and (ii) follows by a similar argument. $\square$

Lemma 6.2. *Suppose that X and $\hat{X}$ are random variables such that $X \in \bigcap_{p\in[1,\infty)} L_p$, and suppose there exists $C = (C_p)_{p\in[1,\infty)} \subset (0,\infty)$ such that $\left\|X - \hat{X}\right\|_p \le C_p$ for all $p \in [1,\infty)$. Let $\theta \in (0,1)$. Then we can choose the function $\varphi_\theta^{X,\hat{X}}$ such that $\varphi_\theta^{X,\hat{X}} = \varphi_{C,\theta}^X$, where the function $\varphi_{C,\theta}^X$ is a bump function that decays faster than any polynomial.*

Proof. The triangle inequality gives that $\hat{X} \in L_p$ and

$$\left\|\hat{X}\right\|_p \leq \left\|X - \hat{X}\right\|_p + \|X\|_p \leq C_p + \|X\|_p.$$

Thus by Chebychev's inequality we have for all $\lambda > 0$ that

$$\mathbb{P}(|X| \geq \lambda) \leq \frac{\mathbb{E}|X|^p}{\lambda^p}$$

and

$$\mathbb{P}(|\hat{X}| \geq \lambda) \leq \frac{\mathbb{E}|\hat{X}|^p}{\lambda^p} \leq \frac{(C_p + \|X\|_p)^p}{\lambda^p}.$$

So we have a polynomial tail estimate for X and $\hat{X}$ that depends only on the constants C_p of the L_p-estimates, not directly on $\hat{X}$. This implies that

$$\left[\mathbb{P}(|X| \geq \lambda) \vee \mathbb{P}(|\hat{X}| \geq \lambda)\right]^\theta \leq \inf_{p \in [1,\infty)} \frac{(C_p + \|X\|_p)^{\theta p}}{\lambda^{\theta p}} \wedge 1 =: \varphi_{C,\theta}^{X,0}(\lambda)$$

for $\lambda > 0$. For $\lambda < 0$ we define $\varphi_{C,\theta}^{X,0}(\lambda) := \varphi_{C,\theta}^{X,0}(|\lambda|)$, and $\varphi_{C,\theta}^{X,0}(0) := 1$. The function $\varphi_{C,\theta}^{X,0}$ satisfies the monotonicity properties of a bump function, but is not necessarily strictly positive. However, if we take a bump function ψ and define

$$\varphi_{C,\theta}^X := \varphi_{C,\theta}^{X,0} \vee \psi,$$

then $\varphi_{C,\theta}^X$ is a bump function suitable for the choice of $\varphi_\theta^{X,\hat{X}}$. Since $\varphi_{C,\theta}^{X,0}$ clearly decays faster than any polynomial and we can choose $\psi(\lambda) = e^{-|\lambda|}$, we see that $\varphi_{C,\theta}^X$ also decays faster than any polynomial. $\square$

Let $\mathcal{P}$ be the set of all polynomials from $\mathbb{R}$ to $\mathbb{R}$. Then we have the following:

Theorem 6.3. *Suppose that φ is a bump function that decays faster than any polynomial. Then $\mathcal{P} \subset \mathcal{G}_{p,\varphi}$ for all $p \in [1, \infty)$.*

Proof. Let $p \in [1, \infty)$ and suppose that $g \in \mathcal{P}$. Then by the fundamental theorem of calculus we have for $x > 0$ that

$$g(x) = g(0) + \int_0^x g'(z)\, dz$$

and for $x \leq 0$ that

$$g(x) = g(0) - \int_x^0 g'(z)\, dz.$$

Thus by defining $c = g(0)$ and a signed measure μ such that

$$d\mu(z) = \operatorname{sgn}(z)\frac{g'(z)}{\varphi(z)}\, dz,$$

we have that $|\mu|$ is σ-finite and the representation $g = c + g^\mu$ holds. Now g' also has only polynomial growth, say $|g'(x)| \leq C(1 + |x|^s)$ for

$s \geq 1$. But φ decays faster than any polynomial, so we have $\varphi(x) \leq \tilde{C}|x|^{-p(s+2)} \wedge 1$ and

$$V_{p,\varphi}(g) = \int_{\mathbb{R}} \varphi^{1+\frac{1}{p}} \, d|\mu| = \int_{\mathbb{R}} \varphi^{\frac{1}{p}}(z)|g'(z)| \, dz$$
$$\leq C\tilde{C}^{\frac{1}{p}} \int_{\mathbb{R}} (|z|^{-(s+2)} \wedge 1)(1 + |z|^s) \, dz < \infty,$$

which implies that $g \in \mathcal{G}_{p,\varphi}$. $\qquad\square$

Let us now come back to the SDE (2.1) and summarize our knowledge:

Corollary 6.4. *Let $p \in [1,\infty)$. Suppose that X is the solution of the equation (2.1), X_T has a bounded density, and X_T^π is an approximation of X_T such that*

$$\|X_T - X_T^\pi\|_p \leq C_p \, |\pi|^\gamma$$

for some constants $\gamma > 0$ and $C_p \geq 0$. Then for any $0 < \varepsilon < \gamma$ we have for $\theta = \frac{\varepsilon}{2\gamma - \varepsilon}$, $\varphi_{C,\theta}^{X_T}$ according to Lemma 6.2 and $g \in \mathcal{G}_{p,\varphi_{C,\theta}^{X_T}}$ that

$$\mathbb{E}|g(X_T) - g(X_T^\pi)|^p \leq 3 \cdot 2^p (\sup f_{X_T})^{1-\frac{\varepsilon}{\gamma}} \left(V_{p,\varphi_{C,\theta}^{X_T}}(g) \right)^p C_{1/\theta}^{1-\frac{\varepsilon}{\gamma}} \, |\pi|^{\gamma - \varepsilon}.$$

Especially, $\mathcal{P} \subset \mathcal{G}_{p,\varphi_{C,\theta}^{X_T}}$.

Proof. By Lemma A.2 we have that $X_T \in \bigcap_{p \in [1,\infty)} L_p$, so by Lemma 6.2 we can choose $\varphi_\theta^{X_T, X_T^\pi} = \varphi_{C,\theta}^{X_T}$, where $\varphi_{C,\theta}^{X_T}$ is a bump function with decay faster than any polynomial. Now using Theorem 5.7 we get for any $q \in [1,\infty)$ and $\theta \in (0,1)$ that

$$\mathbb{E}|g(X_T) - g(X_T^\pi)|^p$$
$$\leq 3 \cdot 2^p (\sup f_{X_T})^{\frac{q(1-\theta)}{q+1}} \left(V_{p,\varphi_{C,\theta}^{X_T}}(g) \right)^p C_q^{\frac{q(1-\theta)}{q+1}} \, |\pi|^{\frac{\gamma q(1-\theta)}{q+1}}.$$

Let $0 < \varepsilon < \gamma$. Now choose $q = \frac{2\gamma}{\varepsilon} - 1$ and let $\theta = 1/q$. Note that $q > 1$ since $\varepsilon < \gamma$. Then

$$\frac{q(1-\theta)}{q+1} = \frac{q-1}{q+1} = 1 - \frac{\varepsilon}{\gamma}$$

and thus we get for all $g \in \mathcal{G}_{p,\varphi_{C,\theta}^{X_T}}$ that

$$\mathbb{E}|g(X_T) - g(X_T^\pi)|^p \leq 3 \cdot 2^p (\sup f_{X_T})^{1-\frac{\varepsilon}{\gamma}} \left(V_{p,\varphi_{C,\theta}^{X_T}}(g) \right)^p C_{1/\theta}^{1-\frac{\varepsilon}{\gamma}} \, |\pi|^{\gamma - \varepsilon}.$$

Moreover, by Theorem 6.3 we have that $\mathcal{P} \subset \mathcal{G}_{p,\varphi_{C,\theta}^{X_T}}$. $\qquad\square$

Remark 6.5. In Corollary 6.4 the function $\varphi_\theta^{X_T, X_T^\pi}$ depends on the distribution of X_T^π and is replaced by the uniform bound $\varphi_{C,\theta}^{X_T}$. However, when considering convergence rate we are looking at partitions with small mesh size. Thus if approximating random variables X_T^π corresponding to partitions with large mesh size had heavy tailed distributions, the use of the uniform bound could unnecessarily narrow down

the class of functions. Therefore in such a case it would be better to take more delicate approach and study the result

$$\mathbb{E}|g(X_T) - g(X_T^\pi)|^p \le 3 \cdot 2^p (\sup f_{X_T})^{1-\frac{\varepsilon}{\gamma}} \left(V_{p,\varphi_\theta^{X_T,X_T^\pi}}(g) \right)^p C_{1/\theta}^{1-\frac{\varepsilon}{\gamma}} |\pi|^{\gamma-\varepsilon}.$$

Corollary 6.4 now gives convergence rates for both Euler and Milstein schemes:

Corollary 6.6. *Let $1 \le p < \infty$. Then for $0 < \varepsilon < \frac{1}{2}$, $\theta = \frac{\varepsilon}{1-\varepsilon}$ and $g \in \mathcal{G}_{p,\varphi_{C,\theta}^{X_T}}$ that*

$$\mathbb{E}|g(X_T) - g(X_T^E)|^p \le 3 \cdot 2^p (\sup f_{X_T})^{1-2\varepsilon} \left(V_{p,\varphi_{C,\theta}^{X_T}}(g) \right)^p C_{1/\theta}^{1-2\varepsilon} |\pi|^{\frac{1}{2}-\varepsilon}.$$

and similarly for $0 < \varepsilon < 1$ and $\theta = \frac{\varepsilon}{2-\varepsilon}$ we have that

$$\mathbb{E}|g(X_T) - g(X_T^M)|^p \le 3 \cdot 2^p (\sup f_{X_T})^{1-\varepsilon} \left(V_{p,\varphi_{C,\theta}^{X_T}}(g) \right)^p C_{1/\theta}^{1-\varepsilon} |\pi|^{1-\varepsilon}.$$

Especially, the statements hold for any $g \in \mathcal{P}$.

Example 6.7. Let us generate a jump function by choosing the measure μ to be a sum of Dirac measures,

$$\mu = \sum_{k \in \mathbb{Z}} \frac{\alpha_k}{\varphi_{C,\theta}^{X_T}(a_k)} \delta_{a_k},$$

where $\alpha_k, a_k \in \mathbb{R}$ for all $k \in \mathbb{Z}$ and $a_k \ne a_l$ for $k \ne l$. Then from the integrability condition for μ we see that $g \in \mathcal{G}_{p,\varphi_{C,\theta}^{X_T}}$ if

$$(6.1) \qquad \sum_{k \in \mathbb{Z}} |\alpha_k| \varphi_{C,\theta}^{X_T}(a_k)^{\frac{1}{p}} < \infty.$$

Therefore the result of Corollary 6.4 holds for jump functions with jumps controlled by the decay of the function $\varphi_{C,\theta}^{X_T}$ in a way that the condition (6.1) is satisfied.

6.1. **Euler scheme.** In the case of the Euler scheme we can again use our knowledge about constants to get more explicit results for the decay of the function $\varphi_\theta^{X_T,X_T^E}$. Let us recall the following result from Bouleau and Lépingle [1]:

Lemma 6.8 ([1, Ch. 5, Lemma B.1.2.]). *For $1 \le p < \infty$,*

$$\sup_{t \le T} |X_t^E| \in L^p$$

and there exist $M(x_0, T, C_T) > 0$ such that

$$\left\| \sup_{t \le T} |X_t^E| \right\|_p \le e^{Mp^2}.$$

Proof. By Theorem A.1 and Lemma A.2 in the Appendix we get

$$\left\|\sup_{t\leq T}|X_t^E|\right\|_p \leq \left\|\sup_{t\leq T}|X_t - X_t^E|\right\|_p + \left\|\sup_{t\leq T}|X_t - x_0|\right\|_p + |x_0|$$

$$\leq (2\sqrt{T} + |x_0|)e^{Mp^2},$$

and we absorb the constant $(2\sqrt{T} + |x_0|)$ into the constant M. □

Theorem 6.9. *We can choose the function $\varphi_\theta^{X_T, X_T^E}$ in a way that $\varphi_\theta^{X_T, X_T^E} \leq \varphi_{\theta,E}^{X_T}$, where $\varphi_{\theta,E}^{X_T}$ is a bump function such that*

(i) *if the functions σ and b are bounded, i.e. $|\sigma|, |b| \leq M$, we have*

$$\varphi_{\theta,E}^{X_T}(z) = \begin{cases} e^{-\frac{\theta(z-(x_0+MT))^2}{2M^2T}} & \text{if } z > \max(x_0 + MT, 0), \\ e^{-\frac{\theta(z-(x_0-MT))^2}{2M^2T}} & \text{if } z < \min(x_0 - MT, 0), \\ 1 & \text{elsewhere,} \end{cases}$$

(ii) *if the functions σ and b are Lipschitz, then there exists $z_0 > 1$ such that we have*

$$\varphi_{\theta,E}^{X_T}(z) = \begin{cases} |z|^{-\frac{2\theta}{3\sqrt{3M}}(\log|z|)^{1/2}} & \text{if } |z| > z_0, \\ 1 & \text{if } |z| \leq z_0, \end{cases}$$

where $M = M(x_0, T, C_T) > 0$.

Proof. (i) We consider the Euler approximation with n time nodes in the integral form (2.2). If we denote

$$L_u := \sum_{k=0}^{n-1} \sigma(t_k, X_{t_k}^E)\chi_{[t_k, t_{k+1})}(u),$$

then by the boundedness of σ and the Novikov condition

$$M_t := e^{\alpha \int_0^t L_u \, dW_u - \frac{\alpha^2}{2}\int_0^t L_u^2 \, du}$$

is a martingale for any $\alpha > 0$, and $\mathbb{E}M_t = 1$. Thus by Chebychev's inequality we have for $\lambda > 1$ that

$$\mathbb{P}\left(e^{\alpha \int_0^T L_u \, dW_u - \frac{\alpha^2}{2}\int_0^T L_u^2 \, du} \geq \lambda\right) \leq \frac{1}{\lambda}.$$

By taking logarithm this implies

$$\mathbb{P}\left(\alpha \int_0^T L_u \, dW_u - \frac{\alpha^2}{2}\int_0^T L_u^2 \, du \geq \lambda\right) \leq e^{-\lambda}$$

for $\lambda > 0$. Since

$$\int_0^T L_u^2 \, du \leq M^2 T,$$

we get

$$\mathbb{P}\left(\int_0^T L_u \, dW_u \geq \frac{\lambda}{\alpha} + \frac{\alpha M^2 T}{2}\right) \leq e^{-\lambda},$$

which we can reparametrize to get

$$\mathbb{P}\left(\int_0^T L_u \, dW_u \geq \lambda\right) \leq e^{\frac{\alpha^2 M^2 T}{2} - \lambda\alpha}$$

for $\lambda > \alpha M^2 T / 2$. Now we can choose $\alpha = \lambda/(M^2 T)$ to get

$$\mathbb{P}\left(\int_0^T L_u \, dW_u \geq \lambda\right) \leq e^{-\frac{\lambda^2}{2M^2 T}}$$

for $\lambda > 0$. A similar proof with $\widetilde{L}_u = -L_u$ shows that

$$\mathbb{P}\left(\int_0^T L_u \, dW_u \leq \lambda\right) \leq e^{-\frac{\lambda^2}{2M^2 T}}$$

for $\lambda < 0$. Therefore, for $\lambda > \max(x_0 + MT, 0)$

$$\begin{aligned}
\mathbb{P}\left(X_T^E \geq \lambda\right) &\leq \mathbb{P}\left(x_0 + \int_0^T L_u \, dW_u + MT \geq \lambda\right) \\
&\leq e^{-\frac{(\lambda - (x_0 + MT))^2}{2M^2 T}},
\end{aligned}$$

and for $\lambda < \min(x_0 - MT, 0)$

$$\begin{aligned}
\mathbb{P}\left(X_T^E \leq \lambda\right) &\leq \mathbb{P}\left(x_0 + \int_0^T L_u \, dW_u - MT \leq \lambda\right) \\
&\leq e^{-\frac{(\lambda - (x_0 - MT))^2}{2M^2 T}}.
\end{aligned}$$

Obviously a similar proof works for the random variable X_T instead of X_T^E, so by the definition of $\varphi_\theta^{X_T, X_T^E}$ the assertion follows. Moreover, to get a bump function we choose the upper bound to be one on the interval $[\min(x_0 - MT, 0), \max(x_0 + MT, 0)]$.

(ii) If σ and b are Lipschitz, then we know from Lemma 6.8 that

$$\left\|X_T^E\right\|_p \leq e^{Mp^2},$$

where the constant $M > 0$ depends on x_0, T and C_T. Now by Chebychev's inequality we have for $\lambda > 0$ that

$$\mathbb{P}(|X_T^E| \geq \lambda) \leq \frac{\mathbb{E}|X_T^E|^p}{\lambda^p} \leq \frac{e^{Mp^3}}{\lambda^p}.$$

Choose $3Mp^2 = \log\lambda$ for $\lambda > \lambda_0 = e^{3M}$. This gives

$$p = \frac{(\log\lambda)^{1/2}}{(3M)^{1/2}},$$

and thus for $\lambda > \lambda_0$ we get

$$\mathbb{P}(|X_T^E| \geq \lambda) = \frac{e^{\frac{1}{3}p\log\lambda}}{\lambda^p} = \lambda^{-\frac{2}{3}p} = \lambda^{-\frac{2}{3\sqrt{3M}}(\log\lambda)^{1/2}}.$$

Again the same proof works for the term $\mathbb{P}(|X_T| \geq \lambda)$ because of Lemma A.2 in the Appendix. $\square$

Theorem 6.10. *Let $c > 0$. If the functions σ and b are bounded, and*

$$g(z) = e^{c|z|^\gamma} \quad if \ \ 0 < \gamma < 2,$$

or

$$g(z) = e^{c|z|^2} \ with \ c < \theta/p,$$

then $g \in \mathcal{G}_{p,\varphi^{X_T}_{\theta,E}}$.

Proof. Since g is not differentiable at zero, define $\tilde{g}(x) := g'(x)$ if $x \neq 0$ and $\tilde{g}(0) := 0$. By choosing a signed measure

$$d\mu(z) = \frac{\mathrm{sgn}(z)\tilde{g}(z)}{\varphi^{X_T}_{\theta,E}(z)}dz,$$

we get that $|\mu|$ is σ-finite and the representation $g = g(0) + g^\mu$ holds. The definition of the class $\mathcal{G}_{p,\varphi^{X_T}_{\theta,E}}$ gives the condition

$$\int_\mathbb{R} \left(\varphi^{X_T}_{\theta,E}(z)\right)^{\frac{1}{p}} |\tilde{g}(z)| \, dz = \int_\mathbb{R} \left(\varphi^{X_T}_{\theta,E}(z)\right)^{\frac{1}{p}} e^{c|z|^\gamma} c\gamma |z|^{\gamma-1} \, dz < \infty,$$

which is by Theorem 6.9 satisfied, because the singularity at zero for $0 < \gamma < 1$ is not too strong, and integrability is determined by the parameters γ, c, θ and p as proposed in the formulation of this Theorem. Similarly we see that the local integrability condition is satisfied. $\square$

7. Lower bound

In this section we find a solution X_1 (i.e. $T = 1$) of an SDE of the type (2.1) such that it gives a lower bound for the approximation rate of the Euler scheme in Theorem 3.7. This is achieved by choosing $X_t = S_t$, the geometric Brownian motion. Let $S_t = e^{W_t - t/2}$ for $t \in [0, 1]$, so that S is a solution of

$$S_t = 1 + \int_0^t S_s \, dW_s$$

and let $U^n := S^E - S$, where S^E is the Euler scheme as defined in (2.2) corresponding to the equidistant partition of $[0, 1]$, i.e. $\pi = (i/n)_{i=0}^n$.

Lemma 7.1. *We have $(W, \sqrt{n}U^n) \Longrightarrow (W, U)$ in the Skorohod topology, where U is the strong solution of the equation*

$$(7.1) \qquad U_t = \int_0^t U_s \, dW_s - \frac{1}{\sqrt{2}} \int_0^t S_s \, dB_s$$

and B is a standard Brownian motion independent of W.

Proof. The statement is an immediate consequence of a result by Jacod and Protter, [10, Corollary 5.4]. $\square$

Theorem 7.2. *There exists $K_0 > 0$ such that*

$$\liminf_{n\to\infty} \sqrt{n} \sup_{K \geq K_0} \mathbb{E}|\chi_{[K,\infty)}(S_1) - \chi_{[K,\infty)}(S_1^E)| > 0,$$

where S_1^E is the equidistant Euler approximation of S_1.

Remark 7.3. Theorem 7.2 states that the convergence rate $\frac{1}{2} - \varepsilon$ for the Euler scheme obtained in Theorem 3.7 and consequently in Theorem 4.5 and Corollary 6.6 is optimal up to the factor ε, i.e. any rate $\gamma > \frac{1}{2}$ leads to a contradiction with the statement of Theorem 7.2.

Proof of Theorem 7.2. Let us consider the setting of Lemma 7.1 and the process U defined by the equation (7.1). If $U_1 = 0$ a.s., then for all $t \in [0, 1]$ we have $U_t = 0$ a.s., which leads to a contradiction. Therefore $\mathbb{P}(U_1 > 0) > 0$ or $\mathbb{P}(U_1 < 0) > 0$. If $\mathbb{P}(U_1 > 0) > 0$, then there exist $\varepsilon \in (0, 1]$, $\delta > 0$ and $K \geq 1 + K_0$ with $K_0 > 0$ such that

$$\mathbb{P}(S_1 \in [K - 1, K), \ U_1 > \varepsilon) = \delta.$$

The case $\mathbb{P}(U_1 < 0) > 0$ can be treated in a similar way by changing the condition $U_1 > \varepsilon$ to $U_1 < -\varepsilon$. By Lemma 7.1 we know that $(W, \sqrt{n}U^n) \Rightarrow (W, U)$ in the Skorohod topology. This implies that $(W_1, \sqrt{n}U_1^n) \Rightarrow (W_1, U_1)$, since the projection mapping π_1, i.e. the mapping $\alpha \mapsto \alpha(1)$ for a process α, is continuous in the Skorohod topology. Because the function $e^{x - \frac{t}{2}}$ is continuous, we have $(S_1, \sqrt{n}U_1^n) \Rightarrow (S_1, U_1)$. Therefore

$$\liminf_{n \to \infty} \mathbb{P}\left(S_1 \in [K - 1, K), \sqrt{n}[S_1^E - S_1] > \varepsilon\right)$$
$$= \liminf_{n \to \infty} \mathbb{P}\left(S_1 \in (K - 1, K), \sqrt{n}\,U_1^n > \varepsilon\right)$$
$$\geq \mathbb{P}\left(S_1 \in (K - 1, K), U_1 > \varepsilon\right)$$
$$= \mathbb{P}\left(S_1 \in [K - 1, K), U_1 > \varepsilon\right),$$

and we see that there exists $n_0 \geq 1$ such that for all $n \geq n_0$

$$\mathbb{P}\left(S_1 \in [K - 1, K), \ [S_1^E - S_1] > \frac{\varepsilon}{\sqrt{n}}\right) \geq \frac{\delta}{2}.$$

Assume a partition $K - 1 = K_0^m < K_1^m < \cdots < K_m^m = K$. Then

$$\sup_{l=1,\ldots,m} \mathbb{P}\left(S_1 \in [K_{l-1}^m, K_l^m), \ [S_1^E - S_1] > \frac{\varepsilon}{\sqrt{n}}\right) \geq \frac{\delta}{2m}.$$

Now choose the partition $(K_l^m)_{l=1}^m$ to be equidistant with

(7.2) $$\frac{1}{m} \leq \frac{\varepsilon}{\sqrt{n}}.$$

Then there exists $l_0 \in \{1, \ldots, m\}$ such that

$$\frac{\delta}{2m} \leq \mathbb{P}\left(S_1 \in [K_{l_0-1}^m, K_{l_0}^m), \ S_1^E > S_1 + \frac{\varepsilon}{\sqrt{n}}\right)$$
$$\leq \mathbb{P}\left(S_1 < K_{l_0}^m, \ S_1^E \geq K_{l_0}^m\right).$$

Let $m = \lceil \sqrt{n}/\varepsilon \rceil$, which satisfies the condition (7.2) for the mesh size. Hence

$$\frac{\delta}{2\lceil \sqrt{n}/\varepsilon \rceil} \ \leq \ \mathbb{P}\left(S_1 < K_{l_0}^m, \ S_1^E \geq K_{l_0}^m\right)$$

$$\leq \quad \mathbb{E}|\chi_{[K^m_{l_0},\infty)}(S_1) - \chi_{[K^m_{l_0},\infty)}(S_1^E)|.$$

Since $\lceil \sqrt{n}/\varepsilon \rceil \leq 2\sqrt{n}/\varepsilon$ we have

$$\mathbb{E}|\chi_{[K^m_{l_0},\infty)}(S_1) - \chi_{[K^m_{l_0},\infty)}(S_1^E)| \geq \frac{\delta}{2\lceil \sqrt{n}/\varepsilon \rceil} \geq \frac{\delta\varepsilon}{4\sqrt{n}}.$$

Therefore

$$\sqrt{n} \sup_{K \geq K_0} \mathbb{E}|\chi_{[K,\infty)}(S_1) - \chi_{[K,\infty)}(S_1^E)| \geq \frac{\delta\varepsilon}{4}$$

for all $n \geq n_0$, which implies the assertion. $\qquad\qquad\square$

APPENDIX A.

Here we prove the following Theorem from the book of Bouleau and Lépingle, [1, pp. 275-276]. The proof is given in the book, but without computing the constant explicitly.

Theorem A.1. *If the assumptions (i)-(iii) in section 2 hold, and $1 \leq p < \infty$, then*

$$\left\| \sup_{0 \leq t \leq T} |X_t - X_t^E| \right\|_p \leq e^{Mp^2} |\pi|^{\frac{1}{2}},$$

where the constant $M > 0$ depends at most on x_0, T and C_T.

For the proof we need the following Lemma:

Lemma A.2. *For $1 \leq p < \infty$ and $0 \leq s \leq t \leq T$,*

$$\left\| \sup_{u \in [s,t]} |X_u - X_s| \right\|_p \leq \sqrt{t - s}\, e^{Mp^2},$$

where $M > 0$ depends at most on x_0, T and C_T.

Proof. Without loss of generality we can suppose that $p \geq 2$. For fixed $s \in [0, T]$ define

$$F(t) = \left\| \sup_{u \in [s,t]} |X_u - X_s| \right\|_p$$

for all $t \in [s, T]$. Then by the Burkholder-Davis-Gundy inequality (as in [1, p. 269]) and the linear growth condition we get

$$\begin{aligned}
F(t) \quad \leq \quad & 2\left\| \int_s^t \sigma(u, X_u)\, dW_u \right\|_p + \left\| \int_s^t |b(u, X_u)|\, du \right\|_p \\
\leq \quad & 8p\left\| \left(\int_s^t |\sigma(u, X_u)|^2\, du \right)^{\frac{1}{2}} \right\|_p + \sqrt{t - s}\left\| \left(\int_s^t |b(u, X_u)|^2\, du \right)^{\frac{1}{2}} \right\|_p \\
\leq \quad & C_T(8p + \sqrt{t - s})\left\| \left(\int_s^t (1 + |X_u|)^2\, du \right)^{\frac{1}{2}} \right\|_p
\end{aligned}$$

$$\leq \ \widetilde{C}\left[\left\|\left(\int_s^t (1+|X_s|)^2\,du\right)^{\frac{1}{2}}\right\|_p + \left\|\left(\int_s^t (|X_u - X_s|)^2\,du\right)^{\frac{1}{2}}\right\|_p\right]$$

$$\leq \ \widetilde{C}\left[\sqrt{t-s}\,\|1+|X_s|\|_p + \left\|\left(\int_s^t \sup_{v\in[s,u]}|X_v - X_s|^2\,du\right)^{\frac{1}{2}}\right\|_p\right]$$

$$\leq \ \widetilde{C}\left[\sqrt{t-s}\,(1+\|X_s\|_p) + \left(\int_s^t \left\|\sup_{v\in[s,u]}|X_v - X_s|\right\|_p^2\,du\right)^{\frac{1}{2}}\right]$$

$$= \ \widetilde{C}\sqrt{t-s}\,(1+\|X_s\|_p) + \widetilde{C}\left(\int_s^t F(u)^2\,du\right)^{\frac{1}{2}},$$

where $\widetilde{C} = C_T(8p + \sqrt{t-s})$. Thus we have

$$\begin{aligned}
F(t)^2 \ &\leq \ 2C_T^2(8p+\sqrt{t-s})^2(t-s)(1+\|X_s\|_p)^2 \\
&\quad + 2C_T^2(8p+\sqrt{t-s})^2\int_s^t F(u)^2\,du
\end{aligned}$$

and by the Gronwall lemma we get

$$F(t)^2 \leq 2C_T^2(8p+\sqrt{t-s})^2(t-s)(1+\|X_s\|_p)^2 e^{2C_T^2(8p+\sqrt{t-s})^2(t-s)}.$$

By taking the square root and choosing a suitable constant M we get

$$F(t) \leq Mp\sqrt{t-s}\,(1+\|X_s\|_p)e^{Mp^2},$$

where $M = M(T, C_T)$. In particular, the above estimate gives that

$$\|X_s\|_p \leq |x_0| + \|X_s - X_0\|_p \leq |x_0| + Mp\sqrt{s}\,(1+|x_0|)e^{Mp^2},$$

so if we redefine the constant M, we get

$$\left\|\sup_{u\in[s,t]}|X_u - X_s|\right\|_p \leq \sqrt{t-s}\,e^{Mp^2},$$

where $M = M(x_0, T, C_T)$. $\qquad\qquad\square$

Proof of Theorem A.1. Suppose that $p \geq 2$. Now define

$$F(t) = \left\|\sup_{s\leq t}|X_s - X_s^E|\right\|_p.$$

Here X^E is the Euler scheme related to the equidistant partition π, and is defined for continuous time by formula (2.2). Then by the Burkholder-Davis-Gundy inequality ([1, p. 269]) we get

$$F(t) \ \leq \ 8p\left\|\left(\int_0^t \sum_{k=0}^{n-1}|\sigma(u, X_u) - \sigma(t_k, X_{t_k}^E)|^2\chi_{[t_k,t_{k+1})}(u)\,du\right)^{\frac{1}{2}}\right\|_p$$

$$+\sqrt{t}\left\|\left(\int_0^t \sum_{k=0}^{n-1}|b(u, X_u) - b(t_k, X_{t_k}^E)|^2 \chi_{[t_k, t_{k+1})}(u)\, du\right)^{\frac{1}{2}}\right\|_p$$

$$\leq 8p\left(\int_0^t \left(\sum_{k=0}^{n-1}\left\|\sigma(u, X_u) - \sigma(t_k, X_{t_k}^E)\right\|_p \chi_{[t_k, t_{k+1})}(u)\right)^2 du\right)^{\frac{1}{2}}$$

$$+\sqrt{t}\left(\int_0^t \left(\sum_{k=0}^{n-1}\left\|b(u, X_u) - b(t_k, X_{t_k}^E)\right\|_p \chi_{[t_k, t_{k+1})}(u)\right)^2 du\right)^{\frac{1}{2}}.$$

By the conditions in Section 2 we have that

$$|\sigma(u, X_u) - \sigma(t_k, X_{t_k}^E)| \leq |\sigma(u, X_u) - \sigma(u, X_{t_k})|$$
$$+|\sigma(u, X_{t_k}) - \sigma(t_k, X_{t_k})| + |\sigma(t_k, X_{t_k}) - \sigma(t_k, X_{t_k}^E)|$$
$$\leq C_T\left(|X_u - X_{t_k}| + (1 + |X_{t_k}|)|u - t_k|^\alpha + |X_{t_k} - X_{t_k}^E|\right),$$

and a corresponding inequality holds for the function b. Thus by Lemma A.2 we have for $u \geq t_k$ that

$$\left\|\sigma(u, X_u) - \sigma(t_k, X_{t_k}^E)\right\|_p$$
$$\leq C_T\left(\left\|X_u - X_{t_k}\right\|_p + \left\|(1 + |X_{t_k}|)|u - t_k|^\alpha\right\|_p + \left\|X_{t_k} - X_{t_k}^E\right\|_p\right)$$
$$\leq C_T\left(e^{Mp^2}|u - t_k|^{\frac{1}{2}} + \left(1 + |x_0| + \sqrt{t_k}e^{Mp^2}\right)|u - t_k|^\alpha + F(u)\right)$$
$$\leq C_T\left((1 + |x_0| + \sqrt{T})e^{Mp^2}(|\pi|^{\frac{1}{2}} + |\pi|^\alpha) + F(u)\right),$$

and again a corresponding inequality holds for b. Denote $C(x_0, T) = 1 + |x_0| + \sqrt{T}$. Now we can continue our estimate for $F(t)$ to get

$$F(t) \leq C_T(8p + \sqrt{T}) \ \cdot$$

$$\left(\int_0^t \left[\sum_{k=0}^{n-1}\left[C(x_0, T)e^{Mp^2}(|\pi|^{\frac{1}{2}} + |\pi|^\alpha) + F(u)\right]\chi_{[t_k, t_{k+1})}(u)\right]^2 du\right)^{\frac{1}{2}}$$

$$\leq C_T(8p + \sqrt{T})\left(\int_0^t \left(C(x_0, T)e^{Mp^2}(|\pi|^{\frac{1}{2}} + |\pi|^\alpha) + F(u)\right)^2 du\right)^{\frac{1}{2}}$$

$$\leq C_T(8p + \sqrt{T})\left[\sqrt{T}C(x_0, T)e^{Mp^2}(|\pi|^{\frac{1}{2}} + |\pi|^\alpha) + \left[\int_0^t F(u)^2\, du\right]^{\frac{1}{2}}\right]$$

$$\leq e^{M_1p^2}(|\pi|^{\frac{1}{2}} + |\pi|^\alpha) + M_1p\left(\int_0^t F(u)^2\, du\right)^{\frac{1}{2}},$$

where $M_1 = M_1(x_0, T, C_T)$ does not depend on p. Taking a square we get

$$F(t)^2 \leq 2e^{2M_1p^2}(|\pi|^{\frac{1}{2}} + |\pi|^\alpha)^2 + 2M_1^2p^2 \int_0^t F(u)^2\, du,$$

and thus Gronwall's Lemma gives

$$F(t)^2 \leq 2e^{2M_1p^2}(|\pi|^{\frac{1}{2}} + |\pi|^\alpha)^2 e^{2M_1^2p^2T} = 2e^{(2M_1 + 2M_1^2T)p^2}(|\pi|^{\frac{1}{2}} + |\pi|^\alpha)^2.$$

By taking square root on both sides and recalling the assumption $\alpha \geq \frac{1}{2}$ this gives

$$F(t) \leq \sqrt{2}e^{(M_1 + M_1^2T)p^2}(|\pi|^{\frac{1}{2}} + |\pi|^\alpha) \leq e^{M_2p^2}|\pi|^{\frac{1}{2}},$$

where $M_2 = M_2(x_0, T, C_T)$. The case $1 \leq p < 2$ follows from the case $p = 2$ by redefining the constant M_2. $\square$

References

[1] NICOLAS BOULEAU, DOMINIQUE LÉPINGLE, *Numerical Methods for Stochastic Processes*. Wiley, 1994.

[2] VLAD BALLY, DENIS TALAY, *The Law of the Euler Scheme for Stochastic Differential Equations: I. Convergence Rate of the Distribution Function*. Probab. Theory Related Fields 104 (1996), no. 1, 43–60.

[3] VLAD BALLY, DENIS TALAY, *The Law of the Euler Scheme for Stochastic Differential Equations: II. Convergence Rate of the Density*. Monte Carlo Methods Appl. 2 (1996), no. 2, 93–128.

[4] COLIN BENNETT, ROBERT SHARPLEY, *Interpolation of Operators*. Academic Press, 1988.

[5] MARÍA EMILIA CABALLERO, BEGOÑA FERNÁNDEZ, DAVID NUALART, *Estimation of Densities and Applications*. J. Theoret. Probab. 11 (1998), no. 3, 831 – 851.

[6] AVNER FRIEDMAN, *Partial Differential Equations of Parabolic Type*. Prentice-Hall, 1964.

[7] NORBERT HOFMANN, THOMAS MÜLLER-GRONBACH, KLAUS RITTER, *The Optimal Discretization of Stochastic Differential Equations*. J. Complexity 17 (2001), No.1, 117 – 153.

[8] NORBERT HOFMANN, THOMAS MÜLLER-GRONBACH, KLAUS RITTER, *Linear vs. Standard Information for Scalar Stochastic Differential Equations*. J. Complexity 18 (2002), 394 – 414.

[9] NORBERT HOFMANN, THOMAS MÜLLER-GRONBACH, *On the Global Error of Itô-Taylor Schemes for Strong Approximation of Scalar Stochastic Differential Equations*. J. Complexity 20 (2004), 732 – 752.

[10] JEAN JACOD, PHILIP PROTTER, *Asymptotic Error Distributions for the Euler Method for Stochastic Differential Equations*. Ann. Prob. 26 (1998), no. 1, 267–307.

[11] IOANNIS KARATZAS, STEVEN E. SHREVE, *Brownian Motion and Stochastic Calculus, Second Edition*, Springer-Verlag, 1991.

[12] PETER E. KLOEDEN, ECKHARD PLATEN, *Numerical Solutions of Stochastic Differential Equations*. Springer-Verlag, 1992.

[13] THOMAS MÜLLER-GRONBACH, *Strong Approximation of Systems of Stochastic Differential Equations*. Habilitation thesis. Darmstadt, 2002.

[14] Thomas Müller-Gronbach, *The Optimal Uniform Approximation of Systems of Stochastic Differential Equations.* Ann. Appl. Probab. 12 (2002), no.2, 664 – 690.

[15] Thomas Müller-Gronbach, *Optimal Pointwise Approximation of SDEs Based on Brownian Motion at Discrete Points.* Ann. Appl. Probab. 14 (2004), no. 4, 1605–1642.

[16] Kyoung-Sook Moon, Anders Szepessy, Raúl Tempone, Georgios E. Zouraris, *Convergence Rates for Adaptive Weak Approximation of Stochastic Differential Equations.* Stoch. Anal. Appl. 23 (2005), no. 3, 511–558.

[17] David Nualart, *The Malliavin Calculus and Related Topics.* Springer-Verlag, 1995.

[18] Walter Rudin, *Real and Complex Analysis, Second Edition.* McGraw-Hill, 1966, 1974.

Department of Mathematics and Statistics, P.O. Box 35 (MaD), FIN-40014 University of Jyväskylä, Finland

E-mail address: avikain@maths.jyu.fi